



From I to D: Pronominal Discourse Functions Across the Sections of Research Articles

Mónica Chávez Muñoz

English Language Centre, Hong Kong Polytechnic University

Room GH 258, Hong Kong, China

Tel: 852 3444 8438

E-mail: mchavez@polyu.edu.hk

Received: 27-02-2013

Accepted: 04-04-2013

Published: 01-05-2013

doi:10.7575/aiac.ijalel.v.2n.3p.218

URL: <http://dx.doi.org/10.7575/aiac.ijalel.v.2n.3p.218>

Abstract

Academic writing has been recently defined as a social activity in disciplinary communities and cultures. As a result, there has been an increasing interest in the study of self as author, focusing on the way writers represent themselves in texts and how writers interact with readers. Additionally, authorial presence has been analysed across languages and across sections of the research article. This study explores the usage pattern of pronominal discourse functions across sections of 60 research articles in the fields of linguistics, psychology and educational research in English and Spanish. Drawing on a modified version of Tang and John's (1999) taxonomy of pronouns, I explain the frequency and usage of pronouns in each section of the research articles. The analysis of items revealed some differences and similarities in terms of linguistic choices writers make in a specific section of the research article in both languages.

Keywords: research articles, text analysis, personal pronouns, metadiscourse, contrastive linguistics

1. Introduction

In the last four decades, the study of interaction in language has had an increasing focus on written texts (Coulthard, 1977; Nystrand, 1986, 1989; Widdowson, 1982; Bakhtin, 1981; McCarthy, 1993; McCarthy and Carter, 1994; Davies, 1994; Bolivar, 1986, 2001; Thompson and Thetela, 1995; Hoey, 2001). Drawing on this research, Hyland (2005b) has created one of the most complete models for the study of interaction in written texts: the metadiscourse model. This model includes a series of linguistic resources that are divided into two major categories: interactive and interactional. On the interactive side, metadiscourse signals include: *transition markers*, *frame markers*, *endophoric markers*, *evidential and code glosses*. On the interactional side, they include: *hedges*, *boosters*, *attitude markers*, *self-mention*, and *engagement markers* (*reader pronouns*, *personal asides*, *shared knowledge*, *directives and questions*). This range of linguistic features has been researched in different genres, such as: popular science articles (Crismore and Farnsworth, 1990), textbooks (Hyland, 1998), student essays (Crismore et al, 1993; Intaraprawat and Steffensen, 1995; Shaw and Liu, 1998; Thompson, 2001; Hyland 2004, Hyland and Tse, 2004), project reports (Hyland, 2000; 2005c), and research articles (Hyland, 1999, 2005a; Hyland and Tse, 2005; Harwood, 2005a, 2005b, 2005c). According to Hyland (2001b), the study of pronominal signals in academic writing has received special attention from language analysts mainly because their proper use in academic genres is still unclear in terms of its use and frequency in the text (Hyland, 2001b). Hence, the existing literature on the analysis of writer and reader pronouns in academic writing has argued that these features could give an insight into the ways writers present their contributions and ideas and interact with other members of the discourse community (Tang and John, 1999; Hyland, 2001b, 2002), and has a self-promotional effect which markets the research carried out (Harwood, 2005a).

1.1 Taxonomies of pronouns in academic writing

Several taxonomies have been proposed for the classification of writer and reader pronouns in academic writing (Tarone, et al. 1998; Bernhardt, 1985; Vassileva, 1998; Ivanič, 1998; Kuo, 1999; Tang and John, 1999; Hyland, 2001a, 2001b, 2002; Harwood 2003, 2005a, 2005b; Starfield and Ravelli, 2006; Sheldon, 2009; Molino, 2010). Additionally, there has been a growing interest from researchers to investigate the use of pronominal discourse functions (PDFs) across sections of research articles (RAs). For instance, an initial analysis of the occurrence of pronouns in the Introduction and Discussion in only one journal was made by Hestlot's (1982). More recent contributions have more concisely researched the distribution and frequency of pronominal signals across sections of Political Science RAs in English (Harwood, 2006), Biology RAs in English written by native and non-native speakers (Martinez, 2005), Biomedical RAs written by English and Spanish writers (Carciu, 2009) and Business Management RAs written in English as L1, and in English as L2, and in Spanish by Spanish academics (Lores Sanz, 2011). Carciu (2009) and Lores Sanz's (2011) studies in particular, have investigated the influence that culture plays on the construction of a text, since different cultures have different ways of using the language for a purpose. As suggested by Atkinson (2004), contrastive rhetoric studies consider culture, whether major national or ethnic cultures, or smaller cultures such as newspapers editorials, economic reports as an influence on textual structure (see Kim, 2009; Kim and Thompson, 2010).

1.2 Hypotheses

In the light of this, the focus of this study is to analyse the sections of RAs in English and Spanish in the fields of Linguistics, Education and Psychology. According to Swales (1990), the RAs sections IMRD have different functions. The Introductory section of the RA emphasises the presence of the disciplinary community, and this is also where the author evaluates and addresses the expectations of the discourse community (Swales, 1990: 140). The Methodology section states the research process, writers list the procedures done in the laboratory in a way that is in principle intended to allow replication of the method (Swales, 1990: 121). Finally, the Results and Discussion are associated with a “cycle of inside-out direction” that states the results themselves, places these within the established literature and reviews their general significance to give closure to the study’s significance (Swales, 1990: 173). Based on this, my initial predictions about the occurrence, frequency, distribution and type of PDFs across sections in both ERAs and SRAs were as follow:

1. PDFs will occur in **all** sections in both languages
2. There will be a continuum of frequency of PDFs across sections in both ERAs and SRAs, from the highest to the lowest: Methodology, Discussion, Introduction, Results.
3. Specific PDFs are expected to occur in both ERAs and SRAs in specific sections as follows:

Introduction-*Representative, Architect,*

Methodology-*Recounter- Methods,*

Results-*Recounter-Results,*

Discussion- *Interpreter, Originator,*

The roles of the *Guide* and *Opinion holder* were expected to occur in all sections.

An important point in my predictions was that regardless of culture specific rhetorical conventions, it was expected that PDFs would behave similarly in both languages. Thus, I predicted that PDFs would be mainly concentrated in the Methodology, as writers commonly list the procedures of the research process and appear as performers of such actions. The second highest concentration of items was expected to occur in the Discussion section where writers’ opinions, claims, contributions, and discussion and interpretation of the results are put forward, these being rhetorical acts of persuasion about the worth of the researchers’ work. The third section in the continuum of frequency would be the Introduction section, where writers use pronominal signals as they interact with the discourse community in stating goals and giving structure to the text. Finally, the Results section was expected to have the lowest concentration of PDFs, as it is here where writers mainly state their results and comment on them (Brett, 1994; Yang and Allison, 2003).

2. Methodology

2.1 Corpus

A total of 60 RAs in 3 different disciplines (linguistics (LI), education (EDU), psychology (PSY)) were collected for the analysis. The chosen disciplines belong to the soft fields (see Becher and Trowler (2001) for a classification of the academic disciplines into soft and hard fields). The criteria for the selection of texts were as follows:

1. The journals in both languages are from the same academic discipline.
2. The selected RAs from each journal are empirical studies and include the following sections: introduction, methodology, results and discussion/conclusion (IMRD).
3. The RAs were all published between 2005 and 2007.
4. All the texts were written by native speakers or near native speakers of the languages in question. It was not tested whether they were native speakers of English and Spanish; however the fact that the articles have been published in high quality journals suggests that the writers had a high command of the language.
5. The articles have been published in leading journals of each discipline, according to the Impact Factor made by the Thompson Institute of Scientific Information, for English articles; and to the Latindex (Scientific Serial Publications Latin American Index) and the Red de Revistas Científicas de América Latina y el Caribe, España y Portugal (Redalyc) for the Spanish articles. A sample list of the articles selected from each journal can be found in Appendix 1. Each text is referred to by abbreviation, that is LIE.RA or LIS.RA, EDUE.RA or EDUS.RA and PSYE.RA or PSYS.RA mean either linguistics, education or psychology Research Articles in English or Spanish, followed by a number (1-5).

2.2 Procedure for data collection

The selected texts were obtained in PDF format and the texts were then converted into plain text format. In order to avoid the analysis of items not belonging to the main sections of the RAs all footnotes, endnotes, reference lists, acknowledgements, quotations and citations were deleted so that only the introduction, methodology, results, discussion sections of the articles made up a clean corpus. Table 1 shows the total number of words in each section of the RAs.

Table 1. Number of words in each section of the RAs

ERAs	Sections					
	Texts	Words	Introduction	Methodology	Results	Discussion
LI	10	111 492	33 251	27 178	24 293	25 181
EDU						
PSY	10	95 574	17 634	34 947	24 750	16 813
Total	10	70 986	15 098	28 556	6 885	18 387
	30	278 052	65 983	90 681	55 928	60 381
SRAs	Texts	Words	Introduction	Methodology	Results	Discussion
LI	10	76 301	31 791	15 072	19 731	7 906
EDU						
PSY	10	73 546	16 616	19 478	25 809	10 327
	10	55 537	20 384	9 390	11 615	12 418
Total	30	205 384	74 791	40 940	57 155	30 651

The data shows a similar number of words in each of the sections across languages, except in the Methodology and Discussion sections where the ERAs show a higher number of words compared to the SRAs. Normalisation of data had to be carried out per 10 000 words due to the difference of words in each corpus.

2.3 Analytical framework

The search for pronominal signals was made using a comprehensive list of pronominal signals in both languages (I, my, my, myself, me, mi, mi / mis, mio/-os/-a/-as, we, our, us, ourselves, nos, nuestro/-os a/-a/ we, our, us, ourselves, nos, nuestro/-os/- a/-as). For the case of subject pronouns in Spanish, the items had to be tagged by verb endings for the plural forms (*nosotros *amos, *emos, *imos*). The quantitative search of items was made using WordSmith tools (Scott, 2008). The qualitative analysis was carried out with a modified version of Tang and John's (1999) continuum of authorial presence that included a new role: *I as the Interpreter*. This new role in the taxonomy represents the writers' interpretation of the final outcome of the research, where their expertise and contribution to the field is expressed. Writers provide meaning and build knowledge according to the results of their study. They bring out the significance of their contribution to the field.

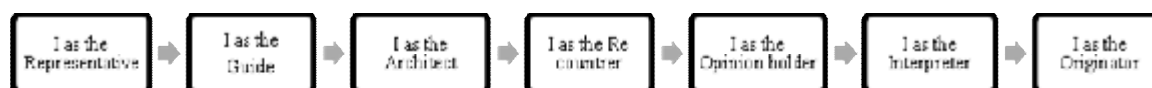


Figure 1. A new continuum of authorial presence in academic writing

3. Results

3.1 Frequency of PDFs

Figure 2 shows the frequency of PDFs across sections in English and Spanish with the normalised data.

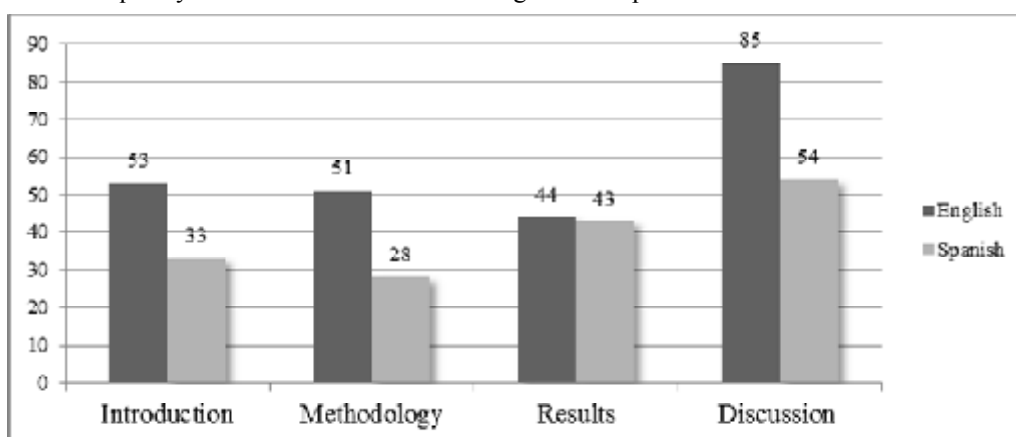


Figure 2. PDFs across sections per 10 000 words

Contrary to my second prediction, the Discussion shows the highest distribution of items in both English and Spanish. The highest distribution of items in the ERAs occurs in the Discussion section, followed by the Introduction, Methodology and finally Results. Similarly to the ERAs, the Discussion section shows the highest distribution of items in the SRAs. However, the Results section is second in the continuum of frequency, followed by Introduction and finally Methodology. Thus, whereas English writers make themselves more present in the text in the Discussion and Introduction, Spanish writers' authorial presence is more frequent in the Discussion and Results section. The latter result partially confirms my initial prediction, that PDFs are the least frequent in the Results sections, at least in the ERAs. This also shows the writers' tendency to be more persuasive at the start and at the close of the articles, and to be more persuasive (and perhaps also more self-promotional) at the start as suggested by Harwood (2005a). However, these results only apply to English, and only partially to Spanish. By heavily using PDFs in the closing section of the RA, writers might be announcing the novelty and value of their work, and it offers them a space to persuade about their work to the reader. One interesting finding is that Spanish writers do not make themselves present in the Methodology section as much as English writers, and in fact, the PDFs are least frequent in the Methodology section of the SRAs. This difference may be explained in terms of the use of passive forms by Spanish writers, who tend prefer to detach themselves from the research process.

As with the overall totals, normalisation of data was carried out in order to make the comparative analysis across sections. Table 2 shows the frequency and distribution of each pronominal discourse function across sections.

Table 2. Distribution of PDFs across sections with normalised data per 10 000 words

	Introduction		Methodology		Results		Discussion	
	ERAs	SRAs	ERAs	SRAs	ERAs	SRAs	ERAs	SRAs
Representative	9.6	5.0	2.4	4.8	3.5	8.8	20.7	12.1
Guide	.15	.80	.88	.97	2.6	2.7	5.4	4.3
Architect	14.3	8.1	5.2	3.6	6.6	12.6	6.9	4.0
Recounter	19.7	13.2	38.2	14.4	23	10.6	34.4	15.1
Opinion holder	4.8	4.1	1.6	1.7	2.3	2.1	0	5.7
Interpreter	0	.13	0	.97	1.6	3.2	4.8	9.4
Originator	3.9	1.4	2.8	.97	4.4	4.9	13.2	3.0

3.2 PDFs in the sections of English and Spanish RAs

3.2.1 PDFs in the Introduction

Contrary to my initial prediction as indicated *RRP* is the most frequent role in both the ERAs (19.7 items per 10 000 words) and SRAs (13.2 items per 10 000 words).

(1) We decided to use diathesis alternations as an approach to characterizing verb behavior, and to use the following verb features to stepwise describe diathesis alternations: (1) syntactic structures, which are relevant for capturing argument functions; (2) prepositions, which are relevant to distinguish, for example, directions from locations; and (3) selectional preferences, which concern participant roles (LIE.RA2)

(2) Primero transcribimos las clases. Debido a que nuestro interés en esta observación está dirigido a las actividades y estrategias que potencia el maestro y no a los detalles del discurso... Utilizamos los signos de puntuación convencionales... (First, we transcribed the classes. Since our interest in this observation is directed to the activities and strategies triggered by the teacher and not by the details of the discourse.... We used the conventional punctuation symbols. (EDUS.RA4)

Examples 1 and 2 occur at the end of the Introduction section before presenting the Methodology section. In both examples writers explain the procedure they followed in their research and announce their methodological procedures and their selected methodological choices, underscoring their work at the start of the RA.

In comparative terms, *RRP* shows a difference in frequency between the ERAs and SRAs (*RRP*-ERAs=19.7 > SRAs=13.2). Perhaps this difference in the use of *RRP* in the Introduction section could be explained if we analyse the small culture in which the genre is embedded. As suggested by Hirano (2009) "smaller discourse communities" in non English speaking contexts tend to be less competitive compared to highly competitive international discourse communities which aim to publish in high impact journals. Hence, the SRAs writers may not promote their methodological steps as often as ERAs writers who flag up their novel methodology in the Introduction more often.

The hypothesis of the relatively frequent occurrence of *R* and *AR* in the Introduction has been confirmed in both the ERAs (*AR*- 14.3 items per 10 000 words; *R*-9.6 items per 10 000 words) and SRAs (*AR*- 8.1 items per 10 000 words; *R*-

5.0 items per 10 000 words). This suggests that the writers in both languages establish a relationship with the discourse community and that they show an awareness of the reader's needs in terms of readability and organisation of the text.

In comparative terms, *R* and *AR* are significantly different between the ERAs and SRAs. These two functions occur almost twice as frequently in the ERAs compared to the SRAs. The greater frequency of *R* suggests that English writers include the reader into the text as a persuasive strategy. This is more commonly found in the English data than the Spanish one. This may be explained in terms of the more competitive environment of international journals, in which gaining the agreement of other fellow members of one's discourse community may be a more difficult task compared to a more "national" and less competitive discourse community.

Also, English writers state goals, signal the overall organisation and section organisation (*prospective architect-73 items*), more often than Spanish writers. This result contradicts Sheldon's (2009) claims about Spanish having a higher frequency of explicit guidelines to the reader, and confirms Valero-Garces's (1996) arguments about the effort readers have to make in order to comprehend a Spanish text due to its implicitness, that is it is the reader's role, at least in this section of the SRAs, to understand the structure of the text. However, these results do not imply that Spanish writers do not signal the organisation of the text, since there is also the option of non reference to writers (Thompson, personal communication, August, 5, 2010 calls this *surrogate say*) to indicate the structure of the paper (Example 3)

(3) El propósito de este artículo es dar lineamientos para la implementación de un programa de lectura extensiva en la enseñanza del inglés como lengua extranjera. El objetivo es promover y mantener el hábito de la lectura por placer, con el fin de que la población estudiantil lea y escriba mejor, enriquezca su vocabulario y logre mejorar también su habilidad comunicativa. (*The purpose of this article is to outline the implementation of an extensive reading programme in teaching English as a foreign language. The objective is to promote and maintain the habit of reading for pleasure, aiming to improve the students writing and reading, enrich their vocabulary and achieve better communicative competence*) (EDUS.RA3)

Items such as example 3 are frequent in the SRAs to signal the organisation of the text. The writers detach themselves from the objective and emphasise the goal of the study itself instead of mentioning themselves as actors of the research process (*our aim, our goal, our objective*, or, more overtly, *we aim to, we wish to etc.*). In this way, they signal the overall organisation and content of the paper without overtly appearing in their text. This may be explained also in terms of the collectivistic tendency of Spanish speaking countries, where the representation of one's self might be considered impolite (Hickey, 2005).

3.2.2 PDFs in the Methodology

Due to the rhetorical function of the Methodology section it was expected that *RRP* would be the most frequently used function, and indeed this was the case in both the ERAs (38.2 items per 10 000 words) and SRAs (14.4 items per 10 000 words). The most frequently used sub-function within *RRP* was *Methods* (ERAs=305 items and SRAs=45 items).

In comparative terms, *RRP* is significantly different when comparing the ERAs and SRAs corpus (*RRP-ERAs=38.3 > SRAs=14.4*). My results correlate with Mur-Duenas' (2007) and Sheldon's (2009) findings when comparing English and Spanish RAs in Business Management and Applied Linguistics and Language Teaching respectively. Both authors found that *RRP* was the most frequent role in the Methodology section, and also that ERAs showed a higher frequency of *RRP* than the SRAs. This result may suggest that English writers are more confident and active in their methodological choices, and claim more responsibility for these choices compared to Spanish writers. This difference may be explained in terms of the rhetorical conventions of Spanish academic discourse, where there might be a preference to create knowledge in a more detached fashion. See example 4.

(4) Para llevar a cabo la cuantificación de las SLS entre las variables textuales: palabras clave-resumen, palabras clave-contenido y resumen-contenido se determinó estadísticamente un corpus de investigación de 22 artículos de investigación científica: 12 de ciencias biológicas y 10 de ciencias sociales. (*In order to carry out the quantification of the SLS between the textual variables: key words, key-content words and content-summary a corpus for investigation made up with 22 research articles was statistically determined: 12 from biological sciences and 10 from social science*). (LIS.RA10)

We can see in example 4 the detachment Spanish writers present in the utterance by using *se determinó* (*was determined*) instead of the active use of a pronoun to signal the writers as performers of the research process, giving objectivity to their research.

The ERAs show *AR* as the second most frequently used function (5.2 items per 10 000 words), *prospective current* being the main pattern to signal the organization of the Methodology section (28 items). This suggests that English

writers also emphasise the organisation of the text in this section of the RA, making this “reader friendly”. In the SRAs, *R* is the second most frequently used function (4.8 items per 10 000 words), *representative researchers* being the most frequent sub-function within *R* to signal the inclusion of the reader.

It is worth pointing out that *R* also varies in frequency between the ERAs and the SRAs (R -ERAs=2.4 < SRAs=4.8). Spanish writers make themselves present in the Methodology as *representative researchers* in the community more often, signaling an inclusion of fellow researchers in the research process. This might be explained in terms of the collectivistic tendency in Hispanic cultures of viewing people as interdependent, leading to an emphasis on group-oriented values (Markus and Kitayama, 1991) and construction of knowledge. See example 5.

(5) Si entendemos que, retomando a Fairclough y Wodak (2000), con el lenguaje se construyen identidades, representaciones y relaciones, queda claro que con el Manual de Carreño se modeló una forma de comportamiento interpersonal y grupal que debía ser interiorizada como norma que el personal Manual prácticamente canonizó. (If we understand that, taking Fairclough and Wodak (2000), language builds identities, representations and relationships, it is clear that with the Manual of Carreno a way of interpersonal and group behaviour was modelled, which should have been internalised as a norm that the Manual canonised. (LIS.RA1)

In example 5 the writer assumes an understanding of field specific concepts from fellow researchers in the discourse community (*si entendemos-if we understand*), building knowledge collectively in this way.

3.2.3 PDFs in the Results

In this section *RRP-Results* was expected to be the most frequently used function. The analysis partly confirms my hypothesis, as *RRP* is the most frequently used function in the ERAs in this section; however *RRP-Methods* (92 items) is the most commonly occurring sub-function, not *RRP-Results* (15 items only). See example 6.

(6) Because of the possible influence of depressive symptoms, comorbid axis I disorders and use of psychotropic drugs on treatment outcome, the regression analysis was repeated by controlling for these variables. We also included baseline BDI scores, comorbid diagnoses (no diagnosis = 0; present diagnosis = 1), and psychotropic drugs during treatment (no drugs = 0; concomitant drugs = 1) as independent variables in the egression analysis. (PSYE.RA8)

Writers in both data sets announce their methodological choices. They present themselves as reliable, careful researchers, whose results can therefore be trusted by claiming responsibility for specific procedures that led to novel results. *RRP* then emphasizes them as having the competence to select appropriate research methods. Contrary to my initial prediction, *AR* is the most frequent function in Results in the SRAs. Thus, the two corpora show different patterns of PDF use in this section. Within this main role *current architect* (41 items) is the most frequently used sub-function. See example 7.

(7) En esta secuencia podemos ver un trabajo deliberado para que los niños apliquen técnicas de lectura para la solución de problemas. Se asegura de contextualizar el problema y que los niños activen esquemas que le permitan ubicar el problema y darle posibles soluciones; les da tiempo para que lean, analicen el problema y lo resuelvan.... Nos interesa aquí resaltar el trabajo que hace de lectura y comprensión de problemas de matemáticas. (In this sequence we can see a deliberate attempt for children to apply reading techniques to solve problems. She makes sure to contextualise the problem and that children activate schemata that allows her to contextualise the problem and give possible solutions; she gives time to read, analyse the problem and solve it.... We are interested here in highlighting the labour of reading and comprehension of mathematical problems. (EDUS.RA)

Items such as example 7 are frequent in the Results section of the SRAs. These usually signal the organisation of the discourse as it unfolds, outlining and explaining the results to the reader. The data may suggest that Spanish writers prefer to use passive forms to signal the results of the research process and build knowledge, instead of the authors' active presence as in English academic discourse, as illustrated in the typical example given below (8)

(8) También se observa un 19,24% de ocurrencias del tiempo futuro absoluto o categórico, sobre todo en la construcción de relaciones causa-efecto: algunas de las

características reseñadas se definen como el origen de manifestaciones atinentes a diverso orden, como se observa en el ejemplo 2: ... *(It is also observed 19.24% of occurrences are future absolute or categoric tense, specially in the construction of cause-effect relations: some of the characteristics are defined as the origin of the manifestations related to a diverse order, as observed in example 2)...* (LIS.RA5)

R, *AR* and *RRP* show significantly different frequency patterns between the ERAs and the SRAs. *R* and *AR* in the SRAs are twice as frequent as in ERAs (*R*-ERAs=3.5 < SRAs= 8.8; *AR*- ERAs=6.6 < SRAs=12.6). However, this differs for *RRP* (*RRP*-ERAs=23 > SRAs=10.6), as it is in the ERAs that this function is twice as frequent as in the SRAs. The latter result confirms Martinez's (2005) findings, with regard to the much higher frequency of *explaining a procedure* (*RRP*) in native compared to non native academic writing in the Results section.

3.2.4 PDFs in the Discussion

Contrary to my initial belief, *OR* and *IN* are not the most frequently used PDFs in the Discussion, but again *RRP* is the most frequently used role in both ERAs and SRAs (34.4 and 15.1 items per 10 000 words respectively). In both ERAs and SRAs *RRP-Methods* is the most frequently used sub-function (118 items and 27 items respectively). See examples 9 and 10.

(9) The fact that giving consent did not depend on the occurrence of an event and that our sample represented the eligible population well speaks against this possibility as a major bias. Our cohort was 78% female and racially homogeneous (white employees), corresponding to Finnish municipal workers in general (39).(PSYE.RA3)

(10) Los menores que hemos entrevistado se ven obligados a trabajar para atender a sus necesidades económicas de subsistencia. Algunos, sobre todo los más pequeños, ayudan en la actividad familiar y no reciben directamente una remuneración, otros trabajan por su cuenta y con el dinero que obtienen ayudan a su familia. *(The under-aged that we have interviewed have been forced to work in order to fulfil their survival economic needs. Some of them, especially the youngest, help in their family's activity and do not receive their earnings directly, some other work on their own and with their earnings, they support their families)* (EDUS.RA7)

This result might be related to the Discussion's "cyclic nature" (Swales, 1990), where writers re-capitulate previous points from the research paper. In self-promotional terms writers continue stating their research steps in this section, perhaps to highlight the relationship between their competent methodological choices and their groundbreaking results and other particularities of their work (Harwood, 2005a).

The second most frequently used function in both the ERAs and SRAs is *R* (20.7 and 12.1 items per 10 000 words respectively). The most common sub-function in the ERAs is *representative researchers* (72 items), whereas the SRAs show *representative everyone* occurring more frequently (21 items). These results could correspond to Move 4-*reference to previous research* (Move 4-*researchers representative*) for the ERAs and Move 7-*deduction* (Move 7-*everyone representative*) for the SRAs. See examples 11 and 12.

(11) Admittedly, this study's findings are limited because they are correlational. However, they fuel previous arguments (e.g., Adams, 1990) that home environment is likely to influence reading performance in various ways, and they extend our understanding of how far in development such effects may be observed. (EDUE.RA8)

(12) A manera de especulación, podría imaginarse que en otro momento de la evolución humana –cuando no dependíamos de la comunicación oral– tal habilidad podría haber sido más relevante en el intercambio social que en el actual, en el que el lenguaje y otros medios de representación del mundo y las emociones nos han desviado de la decodificación del rostro. *(In a speculative way, it can be thought that in another moment of human evolution - when we did not depend on oral communication - such ability could have been more relevant in the social interchange than in the current one, in which language and other means of representation of the world and emotions have diverted us from the decoding of the face).* (PSYS.RA6)

One of the possible explanations for the high occurrence of *R* in both data sets is related to the section's rhetorical function. In the Discussion section writers generally state results and situate these within the wider context in the discourse community and possibly refer back to previous work done in the field.

In comparative terms, *RRP* and *R* show a significant difference in frequency between the ERAs and the SRAs in the Discussion, as they are twice as frequently used in the ERAs ($RRP\text{-ERAs}=34.4 > SRAs=15.1$ and $R\text{-ERAs}=20.7 > SRAs=12.1$). Again, the difference between English and Spanish could be explained in terms of the Spanish writers' preference to detach themselves from their propositions, or perhaps also the issue of "national" versus international communities, in which the latter one demands different rhetorical strategies to persuade the reader in competitive environments.

An interesting finding in this section was the comparative frequency of *IN*. This main role is the third most frequently used function in the SRAs (7.78), whereas the ERAs show a lower frequency of this function ($IN\text{-ERAs}=6.1 < SRAs=9.4$).

(13) ...Como veremos a continuación, esto es algo que se repetía en el grupo de las mujeres del mismo nivel educativo, lo que nos hace suponer que la educación formal y los instrumentos que ésta privilegia favorecen la reflexión sobre la propia práctica. (... *As we will see now, this is something that was repeated in the women group of the same educational level, which makes us suppose that formal education and the instruments that this privileges favour self reflection*) (PSYS.RA10)

In item 13 writers interpret their propositions in an attempt to persuade the reader by generalising claims from their results (Move 8-*hypothesis*).

The most remarkable finding in comparative terms occurs in this section. The *OR* role is four times more frequent in the ERAs compared to the SRAs ($OR\text{-ERAs}=13.2 > SRAs=3.0$). (Example 14)

(14) Our results suggest that school is a powerful contextual tool in understanding reading achievement. Within this framework, the school a child attends is particularly important because of the influences of his peers, who have been affected by their own parents (EDUE.RA8).

This suggests that English writers are more confident when making claims about their outcomes directly by use of the *OR* role. That is, they may market their work and claim responsibility for their novel contribution by presenting themselves as originators of knowledge. On the other hand, Spanish writers are less present and prefer to detach themselves when making claims about their contributions.

4. Discussion & Conclusion

As suggested by Bhatia (2004), genre analysis has now expanded to take account of the social dimension: genre is now envisaged as a strategic space for interaction among the members of a discourse community. This may be related to the high and general occurrence of the *RRP* role in the data set. We could see in the results an interesting pattern of agreement across sections and between corpora: *RRP* was the most frequent main role in all sections in both languages, except for the Results section in Spanish in which *AR* is the most frequent. Another issue is that most of the time the ERAs show a higher frequency of items. However this does not necessarily mean that Spanish writers do not state the steps in the research process (*RRP*), organise the text (*AR*) or show their contributions to the field, but the nature of Spanish academic discourse could be affecting the results. This again could represent the general politeness strategy of Spanish speaking cultures of being detached from propositions (Hickey, 2005), or perhaps the "national" culture in terms of competitiveness for publication has had also an effect on the rhetorical strategies Spanish writers use.

Finally, there are two implications that I would like to point out. First, the occurrence of PDFs has shown that they play a significant role in Linguistics, Education and Psychology RAs. They are on the whole similarly used in both English and Spanish, with some specific and more *locally* oriented differences for Spanish RAs. These findings might be of great value for both novice and professional academic writers in both languages, who often wonder about the correct use of these signals in the text (even in their own native language), fearing to sound too informal. This could be explicitly stated in style manuals and writing in the social sciences courses in both languages, which tend not to agree on pronoun usage and not to include much detail about rhetorical strategies and pronoun usage. Second, although the focus of this study was inter cultural and across sections of the text, and a comparison of the fields was not the main concern, it is worth pointing out that, while a considerable amount of the research carried out in the field has been done in either Linguistics or Education related fields, to my knowledge Psychology has not been explored so much. The function of the *Interpreter* offered rich insights in the Psychology RAs. Thus, I suggest more research in this field in terms of English for Academic Purposes and Specific purposes should be carried out to cover the needs of both novice and professional writers in the field.

Acknowledgements

I am endlessly thankful and indebted to Geoff Thompson for his diligent supervision during my research. His knowledge and support have made the completion of this study possible. I am also grateful to the National Council for Science and Technology of Mexico for funding this study.

References

- Atkinson, D. (2004). Contrasting rhetorics/contrasting cultures: Why contrastive rhetoric needs a better conceptualization of culture. *Journal of English for Academic Purposes*, 3, 277-289.
- Bakhtin, M. (1981). The dialogic imagination: Four essays by M.M. Bakhtin. In M. Holquist, trans. by C. Emerson & M. Holquist. Austin and London: University of Texas Press.
- Becher, T., & Trowler, P. (2001) *Academic tribes and territories: Intellectual enquiry and the culture of disciplines*. Buckingham: SHRE & Open University Press.
- Bernhardt, S. A. (1985). The writer, the reader and the scientific text. *Journal of Technical Writing and Communication*, 15, 163-174.
- Bhatia, V.K. (2004). *Worlds of written discourse: A genre-based view*. London: Continuum.
- Bolivar, A. (1986). *Interaction through written text: a discourse analysis of newspaper editorials*. Unpublished Ph.D. thesis, University of Birmingham.
- Bolivar, A. (2001). The negotiation of evaluation in written text. In M. Scott & G. Thompson, (Eds.), *Patterns of text* (pp. 129-158). Amsterdam: John Benjamins.
- Brett, P. (1994). A genre analysis of the results section of sociology articles. *English for Specific Purposes*, 13, 47-59.
- Carcu, O.M. (2009). An intercultural study of first-person plural references in biomedical writing. *Iberica*, 18, 71-92
- Coulthard, M. (1977). *An introduction to discourse analysis*. London: Longman
- Crismore, A., & Farnsworth, R. (1990). Metadiscourse in popular and professional science discourse. In W. Nash (Ed.), *The writing scholar: Studies in academic discourse* (pp. 118-136). Newbury Park, CA: Sage
- Crismore, A., Markkanen, R., & Steffensen, M.S. (1993). Metadiscourse in persuasive writing. *Written Communication*, 10(1): 39-71.
- Davies, F. (1994). From writer roles to elements of the text: Interactive, organisational and topical. In L. Barbara and M. Scott (Eds.), *Reflections on language learning* (pp. 170-183). Clevedon: Multilingual Matters.
- Harwood, N. (2003). Person markers and interpersonal metadiscourse in academic writing: A multidisciplinary corpus based study of expert and student texts. Unpublished doctoral thesis, Canterbury Christ Church University College, Kent.
- Harwood, N. (2005a). 'Nowhere has anyone attempted... In this article I aim to do just that'. A corpus based study of self-promotional I and we in academic writing across four disciplines. *Journal of Pragmatics*, 37, 1207-1231.
- Harwood, N. (2005b). 'I hoped to counteract the memory problem, but I made no impact whatsoever': Discussing methods in computing science using I. *English for Specific Purposes*, 24, 243-267.
- Harwood, N. (2005c). 'We do not seem to have a theory.... The theory I present here attempts to fill this gap': Inclusive and exclusive pronouns in academic writing. *Applied Linguistics*, 26(3), 343-375.
- Harwood, N. (2006). (In) appropriate personal pronoun use in political science: A qualitative study and a proposed heuristic for future research. *Written Communication*, 23, 324-350.
- Heslot, J. (1982). Tense and other indexical markers in the typology of scientific texts in English. In Hoedt (Ed.), *Pragmatics and LSP* (pp. 83-103). Copenhagen: Copenhagen School of Economics.
- Hickey, L. (2005). Politeness in Spain: Thanks but no thanks. In L. Hickey, & M. Stewart (Eds.), *Politeness in Europe* (pp. 317-330), Clevedon: Multilingual Matters
- Hirano, E. (2009). Research articles introductions in English for specific purposes: A comparison between Brazilian Portuguese and English. *English for Specific Purposes*, 28, 240-250.
- Hopkins, A., & A. Dudley-Evans (1988). A genre-based investigation of the discussion sections in articles and dissertations. *English for Specific Purposes*, 7, 113-22.
- Hoey, M. (2001). *Textual interaction: An introduction to written discourse analysis*. London: Routledge.
- Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics*, 30(4), 437-455.
- Hyland, K. (1999). Disciplinary discourses: Writer stance in research articles. In C.N. Candlin, & K. Hyland, (Eds.), *Writing: Texts, processes and practices* (pp. 122-142). Malaysia: Longman
- Hyland, K. (2000). Options of identity in academic writing. *ELT Journal*, 56 (4), 351- 358.

- Hyland, K. (2001a). Bringing in the reader: addressee features in academic writing. *Written Communication*, 18, 549-574.
- Hyland, K. (2001b). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes*, 20, 207-226.
- Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, 34, 1091-1112.
- Hyland, K. (2004). Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *Journal of Second Language Writing*, 13(2), 133-151.
- Hyland, K. (2005a). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173-192.
- Hyland, K. (2005b). *Metadiscourse: Exploring interaction in writing*. London: Continuum.
- Hyland, K. (2005c). Representing readers in writing: Student and expert practices. *Linguistics and Education*, 16(4), 363-377.
- Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics*, 25(2), 155-177.
- Hyland, K. & Tse, P. (2005). Evaluative *that* constructions. Signaling stance in research abstracts. *Functions of Language*, 12(1), 39-63.
- Intaraprawat, P., & Steffensen M.S. (1995). The use of metadiscourse in good and poor ESL essays. *Journal of Second Language Writing*, 4 (3), 253-272
- Ivanič, R. (1998). *Writing and identity: The discursual construction of identity in academic writing*. Amsterdam: John Benjamins.
- Kim, C. K. (2009). Personal pronouns in English and Korean texts: A corpus-based study in terms of textual interaction. *Journal of Pragmatics*, 41, 2086-2099.
- Kim, C. K. and Thompson, G. (2010). Obligation and reader involvement in English and Korean science popularizations: A corpus-based cross cultural text analysis. *Text & Talk*, 30, 53-73.
- Kuo, C.H. (1999). The use of personal pronouns: Role relationships in scientific journal articles. *English for Specific Purposes*, 18, 121-38.
- Lores Sanz, R. (2011). The construction of the author's voice in academic writing: The interplay of cultural and disciplinary factors. *Text & Talk*, 31, 173-819.
- Lim, J.M.H. (2006). Methods sections of management research articles: A pedagogically motivated qualitative study. *English for Specific Purposes*, 25, 282-309.
- Markus, H.R., & Katayama, S. (1991). Cultures and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224-253
- Martinez, I. A. (2005). Native and non native writers' use of first person pronouns in the different sections of biology research articles in English. *Journal of Second Language Writing*, 14, 174-190.
- McCarthy, M. (1993). Spoken discourse markers in written text.. In J. M. Sinclair, M. Hoey, & G. Fox. (Eds.), *Techniques of Description* (pp. 170-182). London: Routledge
- McCarthy, M., & Carter, R. (1994). *Language as discourse: Perspectives for language teaching*. Harlow: Longman.
- Molino, A. (2010). Personal and impersonal authorial references: A contrastive study of English and Italian Linguistics research articles. *Journal of English for Academic Purposes*, 9, 86-101.
- Mur-Dueñas, P. (2007). 'I/we focus on': A cross-cultural analysis of self-mentions in business management research articles. *Journal of English for Academic Purposes*, 6, 143-162.
- Nwogu, K.N. (1997). The medical research paper: structure and functions. *English for Specific Purposes*, 16, 119-138.
- Nystrand, M. (1986). *The structure of written communication: Studies in reciprocity between writers and readers*. Orlando: Academic Press.
- Nystrand, M. (1989). A social interactive model of writing. *Written Communication*, 6(1), 66-85.
- Scott, M. (2008). *WordSmith Tools*. Oxford: Oxford University Press.
- Starfield, S. & Ravelli, L.J.(2006). 'The writing of this thesis was a process that I could not explore with the positivistic detachment of the classical sociologist': Self and structure in New Humanities research theses. *Journal of English for Academic Purposes*, 5, 222-243.
- Shaw, P., & Liu, T.K. (1998). What develops in the development of second-language writing? *Applied Linguistics*, 19(2), 225-254.
- Sheldon, E. (2009). From one I to another: Discursive construction of self-representation in English and Castilian Spanish research articles. *English for Specific Purposes*, 28, 251-261.

- Swales, J.M. (1990). *Genre Analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Tang, R., & John, S (1999). The I in identity: Exploring writer identity in student academic writing through first person pronoun. *English for Specific Purposes*, 18, 23-39
- Tarone, E., Dwyer, S., Gillet, S., & Icke, V. (1998). On the use of the passive in two astrophysics journal papers. *English for Specific Purposes*, 17, 123-140.
- Thompson, G. (2001). Interaction in academic writing: learning to argue with the reader. *Applied Linguistics*, 22 (1): 58-78.
- Thompson, G. & Thetela, P. (1995). The sound of one hand clapping: The management of interaction in written discourse. *Text*, 15(1): 103-127.
- Valero-Garces, C. (1996). Contrastive ESP rhetoric: Metatext in Spanish-English Economic text. *English for Specific Purposes*, 15, 279-294
- Vassileva, I. (1998). Who am I / who are we in academic writing? A contrastive analysis of authorial presence in English, German, French, Russian and Bulgarian. *International Journal of Applied Linguistics*, 8, 163-90.
- Widdowson, H. (1984). *Explorations in applied linguistics 2*. Oxford: Oxford University Press.
- Yang, R. & Allison, D. (2003). Research articles in applied linguistics: moving from results to conclusions. *English for Specific Purposes*, 22, 365-385.

Appendix I- Sample list of research articles in the data Linguistics

English

LIE .RA1

Ringstetter, C. & Mihov, S. (2005). Orthographic errors in web pages: toward cleaner web corpora. *Computational Linguistics*, 28, 295-340.

LIE .RA2

Schutle im Walde, S. (2006). Experiments on the automatic induction of German semantic verb clauses. *Computational Linguistics*, 32, 159-194.

LIE.RA3

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: evidence against domain specific working memory resources. *Journal of Memory and Language*, 54, 541-553.

LIE.RA4

Engelhardt, P.E., Bailey, K., Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of quality?. *Journal of Memory and Language*, 54, 554-573.

LIE.RA5

Jonsson, M. & Hampton, J. A. (2006). The inverse conjunction fallacy. *Journal of Memory and Language*, 55, 317-334.

Spanish

LIS.RA1

Malaver, I. (2005). Estudio socio pragmático del Manual de Urbanidad y Buenas Maneras de Manuel Antonio Carreño. *Boletín de Lingüística*, 24, 55-77.

LIS.RA2

Domínguez Mujica, C.M. (2005). Marcadores de (in)conclusión en el español hablado en Mérida Venezuela. *Boletín de Lingüística*, 23, 3-22.

LIS.RA3

Chumaceiro Arreaza, I. (2003). El discurso de Hugo Chavez: Bolívar como estrategia para dividir a los venezolanos. *Boletín de Lingüística*, 20: 22-42.

LIS.RA4

Acosta, O. M. (2006). Análisis de introducciones de artículos de investigación publicados en la Revista Núcleo 1985-2003. *Núcleo*, 18, 9-30.

LIS.RA5

Díaz Blanca, L. & Villalobos, A. M. (2006). Los horóscopos: su configuración temporal y moral. *Revista Núcleo*, 18, 31-50.

Education**English****EDUE .RA1**

Zhu, E. (2006). Interaction and cognitive engagement: An analysis of four asynchronous online discussions. *Instructional Science*, 34, 451-480.

EDUE .RA2

Ling Kohl, J. H. (2006). Motivating students of mixed efficacy profiles in technology skills classes: A case study. *Instructional Science*, 34, 423-449.

EDUE .RA3

Levin, T. and Wagner, T. (2006). In their own words: Understanding student conceptions of writing through their spontaneous metaphors in the science classroom. *Instructional Science*, 34, 227-278.

EDUE .RA4

Lièvre, B., Depover, C., Dillenbourg, P. (2006). The relationship between tutoring mode and learners' use of help tools in distance education. *Instructional Science*, 34, 97-127.

EDUE .RA5

Lauer, P., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., Martin-Glenn, M. L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research*, 76, 275-313.

Spanish**EDUS .RA1**

Chacón Araya, Y. and Moncada Jiménez, J. (2006). Relación entre personalidad y creatividad en estudiantes de educación física. *Actualidades Investigativas en Educación*, 6, 1-19.

EDUS .RA2

Carmona Mendoza, R., Crawford Daniels, S., Rivera Calvo, A.C., Zamora González, L. (2006). La aprobación de la lengua escrita en el ciclo de transición con base en la filosofía del lenguaje integral. *Actualidades Investigativas en Educación*, 6, 1-22.

EDUS .RA3

Hernández Herrero, A. (2007). La lectura extensiva: un medio para mejorar la habilidad lingüística de la población estudiantil. *Actualidades Investigativas en Educación*, 7, 1-34.

EDUS .RA4

Cárdenas, M.A., and Rivera, J.F. (2006). El análisis del discurso en el aula: una herramienta para la reflexión. *EDUCERE*, 10, 43-48.

EDUS .RA5

MorLIEs, O.A., Rincón, A.G., Tona Romero, J. (2006). La promoción de la lectura en contextos no escolares y sus implicaciones pedagógicas: estudio exploratorio en Mérida, Venezuela. *EDUCERE*, 10, 283-292.

Psychology**English****PSYE.RA1**

Jacobs, N., Rijdsdijk, F., Derom, C., Vlietinck, R., Delespaul, P., Van Os, J., & Myin-Germeys, I. (2006). Genes making one feel blue in the flow of daily life: A momentary assessment study of gene-stress interaction. *Psychosomatic Medicine*, 68, 201-206.

PSYE.RA2

Berlin, A. A., Kop, W.J., & Deuster, P.A. (2006). Depressive mood symptoms and fatigue after exercise withdrawal: The potential role of decreased fitness. *Psychosomatic Medicine*, 68, 224-230.

PSYE.RA3

Vahtera, J., Kivimäki, M., Väänänen, A., Linna, A., Pentti, J., Helenius, H., & Elovainio M. (2006). Sex Differences in Health Effects of Family Death or Illness: Are Women More Vulnerable Than Men?. *Psychosomatic Medicine*, 68, 283-291.

PSYE.RA4

Johnson-Laird, P. N., Mancini, F., & Gangemi, A. (2006). A hyper-emotion theory of psychological illnesses. *Psychological Review*, 113, 822-841.

PSYE.RA5

Ruini, C., Belaise, C., Brombin C., Caffo, E., & Fava, G. (2006). Well-being therapy in school settings: A pilot study. *Psychotherapy and Psychosomatics*, 75, 331-336.

Spanish**PSYS.RA1**

Robles Montijo, S. Piña López, J. A., & Moreno Rodríguez, D. (2006). Determinantes del uso inconsistente del condón en mujeres que tienen sexo vaginal, oral y anal. *Anales de Psicología*, 22, 200-204.

PSYS.RA2

López Durán, A. & Becoña Iglesias, E. (2006). El craving en personas dependientes de la cocaína. *Anales de Psicología*, 22, 205-211.

PSYS.RA3

Remor, E., Amoros, M., & Carroblies, J.A. El optimismo y la experiencia de ira en relación con el malestar físico. *Anales de Psicología*, 22, 37-44.

PSYS.RA4

Pardo, C.F. & Burbano Valente, J. (2006). Las trampas de la velocidad: Análisis de la lectura cultural de dos organizaciones. *Universitas Psychological*, 6, 131-142.

PSYS.RA5

Cárdenas, M. (2006). Escala de racismo moderno: Propiedades psicométricas y su relación con variables psicosociales. *Universitas Psychological*, 6, 255-262.